Bivariate Distribution

Let Y be the number of cars sold at Saratoga dealership over 300 days

X be the number of cars sold at Geneva dealership over 300 days

Joint probability

$P(Y=0 \text{ and } X=0) = 0.07$

$P(Y=0, X=1) = 0.07$

$P(Y=2, X=3) = 0.02$

$P(Y=3, X=1) = 0.06$

Marginal probability

$P(Y=0) = 0.18$

$P(Y=2) = 0.24$

$P(X=3) = 0.0867$

$P(X=1) = 0.3700$

$$Var(X) = E[(X - E[X])^2]$$

$$= \sum_i (x_i - E[X])^2 f(x)$$

$$E[X] = \frac{1}{n} \sum_i x_i f(x_i)$$

outcome $\qquad$ $P(X = x)$

$Var(Y) = 1.244$

Compute $E[Y] = \sum_i y_i f(y_i)$

$P(X=x)$

y	$f(y)$	$y f(y)$	$y - E[Y]$	$(y - E[Y])^2$	$(y - E[Y])^2 f(y)$
0	0.18	0	-1.40	1.96	0.3528
1	0.29	0.29	-0.40	0.16	0.0464
2	0.24	0.48	0.60	0.36	0.0864
3	0.14	0.42	1.60	2.56	0.3584
4	0.04	0.16	2.60	6.76	0.2704
5	0.01	0.05	3.60	12.96	0.1296

$\sum y f(y) = 1.40 = E[y]$

$\sum (y - E[Y])^2 f(y)$

Table 5.7 Number of Automobiles Sold at Dicarlo's Saratoga and Geneva Dealerships Over 300 Days

$P(X=0 \text{ and } Y=3)$

Geneva Dealership (X)	Saratoga Dealership (Y) 0	1	2	3	4	5	Total
0	21	30	24	9	2	0	86
1	21	36	33	18	2	1	111
2	9	42	9	12	3	2	77
3	3	9	6	3	5	0	26
Total	54	117	72	42	12	3	300

$X=0, Y=3$

$X=2, Y=1$

Table 5.8 Bivariate Empirical Discrete Probability Distribution for Daily Sales at Dicarlo Dealerships in Saratoga and Geneva, New York

joint probability

marginal probability

X Geneva Dealership	Saratoga Dealership Y						
	0	1	2	3	4	5	Total
0	.0700	.1000	.0800	.0300	.0067	.0000	.2867
1	.0700	.1200	.1100	.0600	.0067	.0033	.3700
2	.0300	.1400	.0300	.0400	.0100	.0067	.2567
3	.0100	.0300	.0200	.0100	.0167	.0000	.0867
Total	.18	.39	.24	.14	.04	.01	1.0000

$= P(X=0)$

$= P(X=1)$

$= P(X=2)$

$= P(X=3)$

$P(Y=0)$ $P(Y=1)$ $P(Y=2)$ $P(Y=5)$

Number of Automobiles Sold at Dicarlo's Saratoga and Geneva Dealerships Over 300 Days

	Saratoga Dealership Y							
X Geneva Dealership	0	1	2	3	4	5	Total	
0	21 f_{11}	30 f_{12}	24 f_{13}	9 f_{14}	2 f_{15}	0 f_{16}	86	r_1
1	21 f_{21}	36	33	18	2	1	111	r_2
2	9 f_{31}	42	9	12	3	2	77	r_3
3	3 f_{41}	9	6	3	5	0 f_{45}	26	r_4
Total	54	117	72	42	12	3	300	n
	c_1	c_2	c_3	c_4	c_5	c_6		

8

Bivariate Empirical Discrete Probability Distribution for Daily Sales at Dicarlo Dealerships in Saratoga and Geneva, New York

	Saratoga Dealership						
Geneva Dealership	0	1	2	3	4	5	Total
0	.0700	.1000	.0800	.0300	.0067	.0000	.2867
1	.0700	.1200	.1100	.0600	.0067	.0033	.3700
2	.0300	.1400	.0300	.0400	.0100	.0067	.2567
3	.0100	.0300	.0200	.0100	.0167	.0000	.0867
Total	.18	.39	.24	.14	.04	.01	1.0000

$P(X=x \cap Y=y)$

$\uparrow$

$P(X=x \text{ and } Y=y)$

$\uparrow$

Bivariate Distribution

1. Joint Probability : $P(X=x, Y=y) = \dfrac{f_{ij}}{n}$

2. Marginal Probability: $P(X=x) = \dfrac{r_j}{n}$

   $P(Y=y) = \dfrac{c_i}{n}$

3. Conditional Probability:

joint

$$\underbrace{P(X=x \mid Y=y)}_{\text{conditional}} = \frac{P(X=x, Y=y)}{\underbrace{P(Y=y)}_{\text{marginal}}} = \frac{f_{ij}}{r_j}$$

$$P(Y=y \mid X=x) = \frac{P(Y=y, X=x)}{P(X=x)} = \frac{f_{ij}}{c_i}$$

4. $E[X] = \sum x P(X=x)$

5. $E[Y] = \sum y P(Y=y)$

Expectation

6. Conditional Expectations (averag value)

$$E[X \mid Y=y] = \sum x P(X=x \mid Y=y)$$

$$E[Y \mid X=x] = \sum y P(Y=y \mid X=x)$$

$x = 0, 1, 2, 3$; $y = 0, 1, 2, 3, 4, 5$

Exercises

find the following probabilities

1) $P(X=2, Y=0) = 9/300 = 0.030$

2) $P(X=0) = 86/300 = 0.287$

3) $P(Y=4) = 12/300 = 0.040$

4) $P(X=1, Y=2) = 33/300 = 0.110$

5) $P(X=1 \mid Y=2) = \dfrac{P(X=1, Y=2)}{P(Y=2)}$

$$= \frac{33/\cancel{300}}{72/\cancel{300}}$$

$$= \frac{33}{72} = 0.458$$

Number of Automobiles Sold at Dicarlo's Saratoga and Geneva Dealerships Over 300 Days

X Geneva Dealership	Saratoga Dealership Y 0	1	2	3	4	5	Total
0	21	30	24	9	2	0	86
1	21	36	33	18	2	1	111
2	9	42	9	12	3	2	77
3	3	9	6	3	5	0	26
Total	54	117	72	42	12	3	300

$$E[X] = \sum x P(X=x)$$

$$= 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2P(X=2) + 3P(X=3)$$

$$= 0 + \frac{111}{300} + 2\left(\frac{77}{300}\right) + 3\left(\frac{26}{300}\right)$$

$$= 1.143$$

6) $P(X=2 \mid Y=2) = 9/72$

7) $P(X=3 \mid Y=2) = 6/72$

8) $P(X=0 \mid Y=2) = 24/72$

9) $P(Y=0 \mid X=3) = 3/26$

10) $P(Y=1 \mid X=3) = 9/26$

11) $P(Y=2 \mid X=3) = 6/26$

12) $P(Y=3 \mid X=3) = 3/26$

13) $P(Y=4 \mid X=3) = 5/26$

14) $P(Y=5 \mid X=3) = 0/26$

Number of Automobiles Sold at Dicarlo's Saratoga and Geneva Dealerships Over 300 Days

(X) Geneva Dealership	Saratoga Dealership (Y) 0	1	2	3	4	5	Total
0	21	30	24	9	2	0	86
1	21	36	33	18	2	1	111
2	9	42	9	12	3	2	77
3	3	9	6	3	5	0	26
Total	54	117	72	42	12	3	300

$$E[Y] = \sum y P(Y=y)$$

$$= 0P(Y=0) + 1P(Y=1) + 2P(Y=2) + 3P(Y=3) + 4P(Y=4) + 5P(Y=5)$$

$$= 0 + \frac{117}{300} + 2\left(\frac{72}{300}\right) + 3\left(\frac{42}{300}\right) + 4\left(\frac{12}{300}\right) + 5\left(\frac{3}{300}\right)$$

$$= 1.500$$

The average number of cars sold by Saratoga dealership given that Geneva dealership sold 3 cars

$$E[Y|X=3] = \sum y P(Y=y|X=3)$$

conditional expectation

$$= 0P(Y=0|X=3) + 1P(Y=1|X=3) + 2P(Y=2|X=3)$$
$$+ 3P(Y=3|X=3) + 4P(Y=4|X=3) + 5P(Y=5|X=3)$$

$$= 0 + 1\left(\frac{9}{26}\right) + 2\left(\frac{6}{26}\right) + 3\left(\frac{3}{26}\right) + 4\left(\frac{5}{26}\right) + 5\left(\frac{0}{26}\right)$$

$$= 1.923$$

Therefore the average number of cars sold by Saratoga dealership given that Geneva dealership sold 3 cars is 1.923

Covariance of Random Variables x and y (See *)

$$\sigma_{xy} = [Var\,(x+y) - Var\,(x) - Var\,(y)]/2 \tag{5.6}$$

Table 5.10 Calculation of the Expected Value and Variance of Daily Automobile Sales at Dicarlo Motors' Geneva Dealership

x	$f(x)$	$xf(x)$	$x - E(x)$	$[(x - E(x)]^2$	$[x - E(x)]^2 f(x)$
0	.2867	.0000	−1.1435	1.3076	.3749
1	.3700	.3700	−.1435	0.0206	.0076
2	.2567	.5134	.8565	0.8565	.1883
3	.0867	.2601	1.8565	1.8565	.2988
		$E(x) = 1.1435$			$Var(x) = .8696$

Table 5.10 **Calculation of the Expected Value and Variance of Daily Automobile Sales at Dicarlo Motors' Geneva Dealership**

x	*f*(*x*)	*xf*(*x*)	*x* − *E*(*x*)	$[(x - E(x)]^2$	$[x - E(x)]^2 f(x)$
0	.2867	.0000	−1.1435	1.3076	.3749
1	.3700	.3700	−.1435	0.0206	.0076
2	.2567	.5134	.8565	0.8565	.1883
3	.0867	.2601	1.8565	1.8565	.2988
		$\boldsymbol{E(x) = 1.1435}$			$\boldsymbol{Var(x) = .8696}$

Correlation between Random Variables *x* and *y*

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \tag{5.7}$$

DiCarlo dealerships. First we compute the standard deviations for sales at the Saratoga and Geneva dealerships by taking the square root of the variance.

$$\sigma_x = \sqrt{.8696} = .9325$$
$$\sigma_y = \sqrt{1.25} = 1.1180$$

Now we can compute the correlation coefficient as a measure of the linear association between the two random variables.

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{.1350}{(.9325)(1.1180)} = .1295$$